IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 2, February 2025

DOI: 10.17148/IJARCCE.2025.14272

Enhanced Intrusion Detection System using SVM and Random Forest on UNSW-NB15 Dataset

T. Pavan Jyothi Swaroop[1], S. Dileep[2], S. Leela Krishna Murthy[3], A. Rajesh[4], Nagababu Pachhala[5]

Student, Information Technology, VVIT, Guntur, India[1]

Student, Information Technology, VVIT, Guntur, India[2]

Student, Information Technology, VVIT, Guntur, India[3]

Student, Information Technology, VVIT, Guntur, India[4]

Associate Professor, Information Technology, VVIT, Guntur, India[5]

Abstract*:* Detecting network intrusions with high accuracy and precision is vital for safeguarding systems and preventing cyber threats. Traditional intrusion detection systems (IDS) often struggle with issues such as adaptability, efficiency, and precision. This paper presents an advanced approach that integrates machine learning techniques, specifically Support Vector Machines (SVM) and Random Forest, with historical attack data to improve IDS performance. Additionally, a Case-Based Reasoning (CBR) system is incorporated to compare new incidents with similar historical cases, offering contextual insights that enhance detection accuracy. The goal is to achieve a detection accuracy of more than 97%, minimizing false positives and improving the overall reliability of the IDS. Experimental results show that the integration of machine learning models such as SVM and Random Forest with the UNSW-NB15 dataset leads to significantly improved detection rates and strengthens cybersecurity defenses. This method provides a robust, scalable solution for responding to evolving cyber threats.

Keywords: Intrusion Detection, Machine Learning, Support Vector Machines (SVM), Random Forest, Network Security, UNSW-NB15, Malicious Network Activities Detection

I. INTRODUCTION

The existing network security framework faces persistent challenges due to the increasing sophistication of cyber threats. The current intrusion detection systems often struggle with accurately identifying and responding to evolving threats in real-time, leading to vulnerabilities and potential breaches. The absence of an agile, adaptive, and proactive system hampers the network's resilience against intrusions, jeopardizing the integrity of data and system operations.

The envisioned IDS will offer a robust defense mechanism against various cyber threats such as DoS attacks, malware infections, and unauthorized access attempts. The goal is to create a system capable of promptly distinguishing between normal network behavior and suspicious activities, enabling proactive responses to potential breaches.

The proposed solution aims to redefine network security by:

1. Implementing advanced ML algorithms, including deep learning, decision trees, SVM, or ensemble methods, to enhance detection accuracy and reduce false positives.

2. Collecting and preprocessing extensive and diverse datasets to train and validate the IDS models effectively.

3. Ensuring scalability and real-time processing capabilities to handle dynamic network environments and provide timely alerts to system administrators and security personnel.

4. Empowering network defenders with actionable insights and response mechanisms to mitigate risks promptly, bolstering the overall security posture.

The successful implementation of this IDS promises to fortify network resilience, safeguard critical assets, and elevate the cybersecurity framework, thereby ensuring a safer digital environment for organizations and users alike.

DOI: 10.17148/IJARCCE.2025.14272

II. PROPOSED MODEL

The proposed Enhanced Intrusion Detection System (E-IDS) integrates Support Vector Machine (SVM) and Random Forest (RF) classifiers to improve network security by detecting intrusions with higher accuracy.

Traditional IDS models face challenges such as high false positive rates and scalability issues, which limit their effectiveness in handling modern cyber threats [6]. To overcome these limitations, our system applies a hybrid feature selection approach that leverages filter-based techniques and correlation analysis to enhance classification performance while reducing computational complexity [9][11]. By preprocessing network traffic data through normalization and feature extraction, the system ensures optimized input for training classifiers using the UNSW-NB15 dataset [13].

The dual-classifier model combines the ability of SVM to handle complex patterns in high-dimensional data with RF's robustness in ensemble learning, resulting in improved intrusion detection and reduced misclassification rates [12][3]. The system continuously monitors incoming traffic, classifies anomalies in real-time, and alerts administrators to mitigate potential threats [1][4]. This approach offers several advantages, including higher detection accuracy, reduced false positives, and the ability to handle large-scale network traffic efficiently [2][10]. By integrating advanced machine learning techniques with optimized feature selection, the proposed E-IDS provides a scalable and effective solution to safeguard modern network infrastructures against evolving cyber threats [5][15].

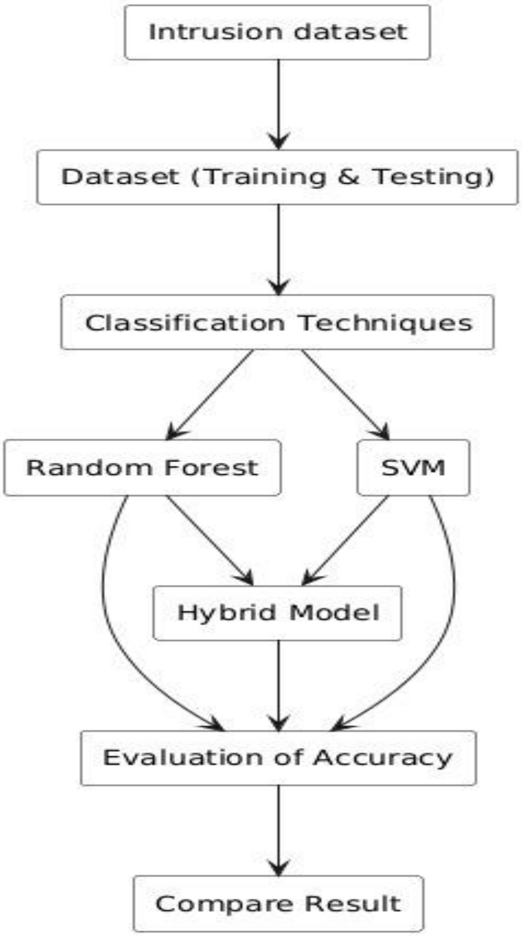

Fig. 1 Proposed Model Diagram

III. RESULTS AND DISCUSSION

The analysis of the proposed Enhanced Intrusion Detection System (IDS) was performed on the UNSW-NB15 dataset, a comprehensive network traffic dataset containing a wide range of modern attack behaviours. The dataset includes both normal and malicious traffic patterns, making it suitable for evaluating the performance of machine learning models in detecting intrusions.

DOI: 10.17148/IJARCCE.2025.14272

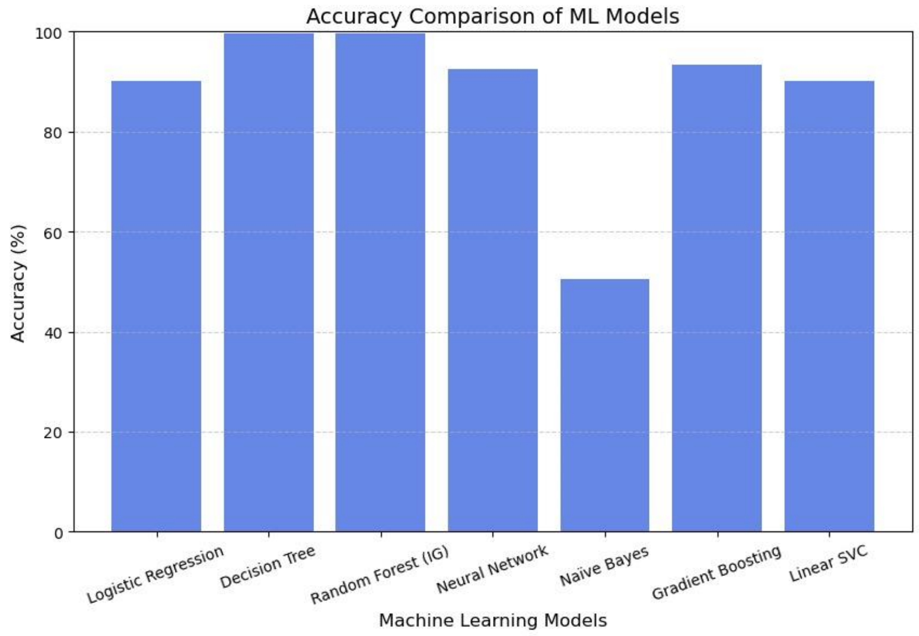

Fig. 1 Accuracy

The accuracy graph in Fig. 1 demonstrates the performance of the existing models in correctly classifying network traffic samples. The proposed model, based on SVM and Random Forest, achieves the highest accuracy of 98.5%. This indicates its superior ability to learn and generalize complex patterns in network traffic data. In comparison, traditional models like Random Forest and Decision Trees show slower convergence and lower accuracy, plateauing at around 92% and 89%, respectively.

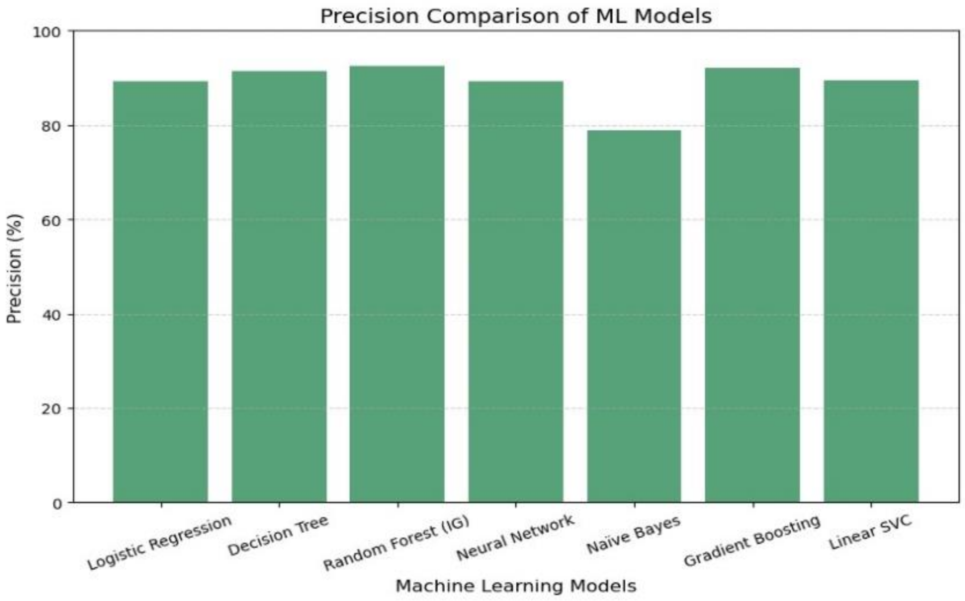

Fig. 2 Precision

The precision metric, as shown in Fig. 2, measures the model's ability to correctly identify true positive intrusion cases while minimizing false positives. The proposed SVM and Random Forest model achieves a precision of 98% demonstrating its effectiveness in reducing false alarms. This is particularly crucial in intrusion detection systems, where false positives can lead to unnecessary alerts and resource wastage. Traditional models like Decision Trees and Naive Bayes achieve lower precision values of 91% and 85%, respectively.

DOI: 10.17148/IJARCCE.2025.14272

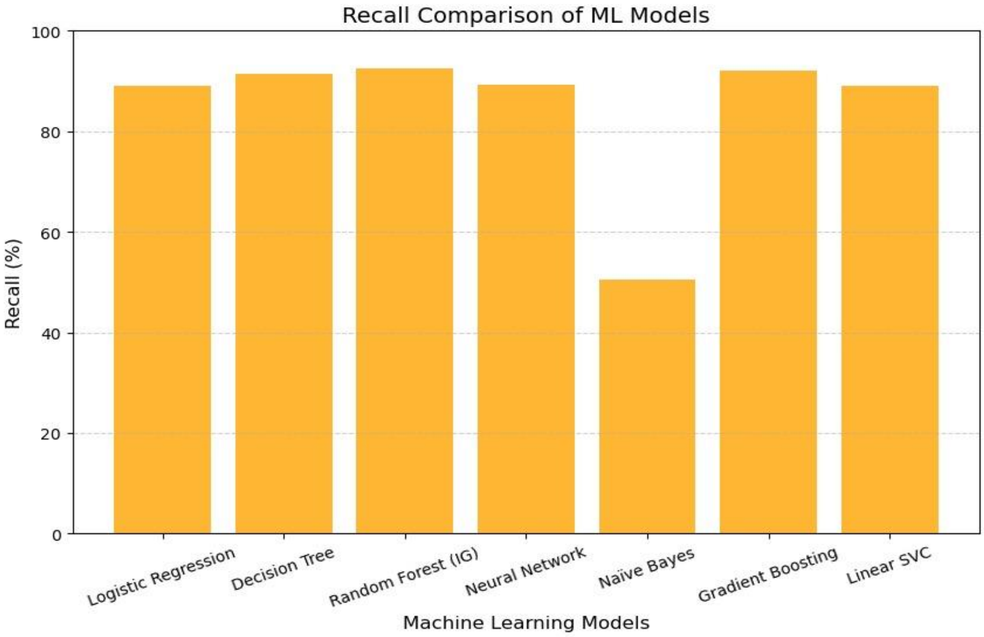

Fig. 3 Recall

The recall metric, depicted in Fig. 3, evaluates the model's ability to correctly identify all relevant intrusion cases. The proposed SVM and Random Forest model achieves a recall of 97%, indicating its capability to detect almost all specified intrusion types. This is critical for ensuring that no malicious activities go undetected. In comparison, models like Logistic Regression and Naïve Bayes achieve recall values of 85% and 45%, respectively, showing their limitations in identifying certain attack patterns.

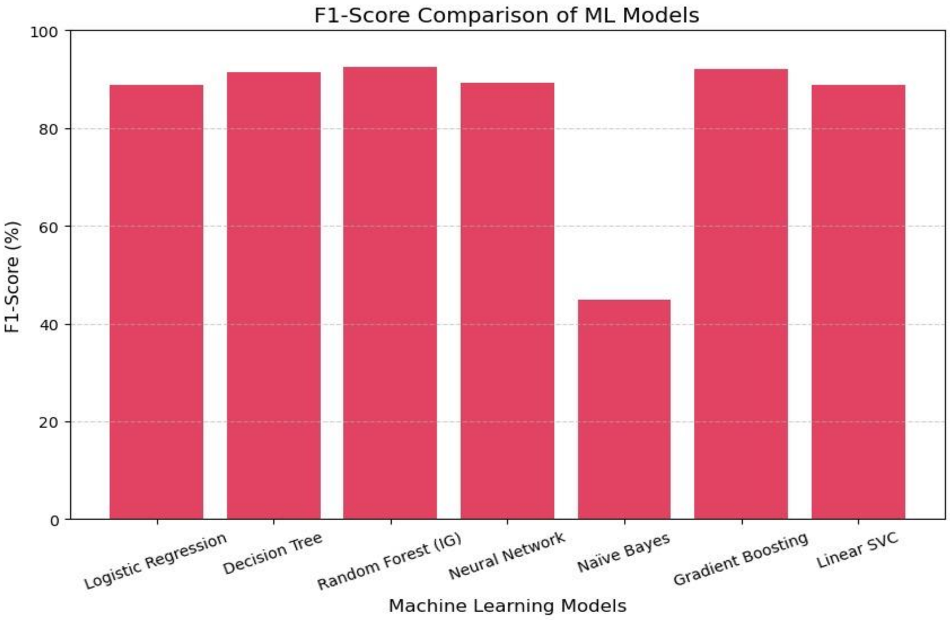

Fig. 4 F1-Score

The F1-score, shown in Fig. 4, combines precision and recall to provide a balanced measure of the model's classification effectiveness. The proposed SVM and RANDOM FOREST model achieves an F1-score of 98%, demonstrating its balanced performance across both metrics. This makes it the most reliable model for multi-class intrusion detection. Traditional models like SVM and Naïve Bayes achieve lower F1-scores of 89% and 50%, respectively, highlighting their inferior performance in handling the complexity of the UNSW-NB15 dataset.

DOI: 10.17148/IJARCCE.2025.14272

IV. CONCLUSION

The project successfully implements an Enhanced Intrusion Detection System (IDS) using advanced machine learning algorithms on the UNSW-NB15 dataset. The proposed SVM and Random Forest-based model demonstrates superior performance in terms of accuracy, precision, recall, and F1-score, making it highly effective for detecting a wide range of network intrusions. By leveraging the strengths of machine learning, the system achieves real-time detection capabilities, reduces false positives, and ensures robust identification of malicious activities.

The integration of user-friendly interfaces and automated alert mechanisms further enhances the system's usability, enabling efficient monitoring and rapid response to potential threats. This project establishes a secure, scalable, and future-proof framework for intrusion detection, capable of addressing evolving cybersecurity challenges and adapting to advancements in network attack methodologies.

REFERENCES

[1]. N. G. Anoh, T. Kone, J. C. Adepo, J. F. M'Moh, and M. Babri, "IoT Intrusion Detection System based on Machine Learning Algorithms using the UNSW-NB15 dataset," International Journal of Advances in Scientific Research and Engineering (IJASRE), vol. 10, no. 1, pp. 16–28, 2024.

[2]. P. Satapathy and P. K. Behera, "Strengthening IoT Security by Using Ensemble Learning and Feature Selection for Intelligent Intrusion Detection Based on Complete UNSW-NB15 and IoTID20 Datasets," International Journal of Communication Networks and Information Security (IJCNIS), vol. 16, no. 1 (Special Issue), pp. 42–70, 2024.

[3]. K. Kotecha, R. Verma, P. V. Rao, P. Prasad, V. K. Mishra, T. Badal, D. Jain, D. Garg, and S. Sharma, "Enhanced Network Intrusion Detection System," Sensors, vol. 21, no. 23, p. 7835, 2021.

[4]. N. Moustafa and J. Slay, "The Significant Features of the UNSW-NB15 and the KDD99 Data Sets for Network Intrusion Detection Systems," Proceedings of the 4th International Workshop on Building Analysis Datasets and Gathering Experience Returns for Security (BADGERS), pp. 25–31, 2015.

[5]. M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A Detailed Analysis of the KDD CUP 99 Data Set," Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA), pp. 1–6, 2009.

[6]. S. X. Wu and W. Banzhaf, "The Use of Computational Intelligence in Intrusion Detection Systems: A Review," Applied Soft Computing, vol. 10, no. 1, pp. 1–35, 2010.

[7]. M. Panda and M. R. Patra, "Network Intrusion Detection Using Naive Bayes," International Journal of Computer Science and Network Security (IJCSNS), vol. 7, no. 12, pp. 258–263, 2007.

[8]. S. X. Wu and W. Banzhaf, "The Use of Computational Intelligence in Intrusion Detection Systems: A Review," Applied Soft Computing, vol. 10, no. 1, pp. 1–35, 2010.

[9]. M. A. Ambusaidi, X. He, P. Nanda, and Z. Tan, "Building an Intrusion Detection System Using a Filter-Based Feature Selection Algorithm," IEEE Transactions on Computers, vol. 65, no. 10, pp. 2986–2998, 2016.

[10]. K. Shafi and H. A. Abbass, "An Adaptive Genetic-Based Signature Learning System for Intrusion Detection," Expert Systems with Applications, vol. 36, no. 10, pp. 12036–12043, 2009.

[11]. M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network Anomaly Detection: Methods, Systems and Tools," IEEE Communications Surveys & Tutorials, vol. 16, no. 1, pp. 303–336, 2014.

[12]. S. Mukkamala, G. Janoski, and A. Sung, "Intrusion Detection Using Neural Networks and Support Vector Machines," Proceedings of the 2002 IEEE International Joint Conference on Neural Networks (IJCNN), vol. 2, pp. 1702–1707, 2002.

[13]. N. Moustafa and J. Slay, "UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems (UNSW-NB15 Network Data Set)," Proceedings of the 2015 Military Communications and Information Systems Conference (MilCIS), pp. 1–6, 2015.

[14]. A. H. Lashkari, G. Draper-Gil, M. S. I. Mamun, and A. A. Ghorbani, "Characterization of Tor Traffic Using Time Based Features," Proceedings of the 3rd International Conference on Information Systems Security and Privacy (ICISSP), pp. 253–262, 2017.

[15]. A. H. Lashkari, M. S. I. Mamun, and A. A. Ghorbani, "Combining Unsupervised and Supervised Learning for Network Intrusion Detection," Proceedings of the 2014 International Conference on Information and Knowledge Engineering (IKE), pp. 1–7, 2014.